ELSEVIER
Journal of Computational and Applied Mathematics 126 (2000) 319–338
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.nl/locate/cam

Improved error estimation of dynamic finite element methods for second-order parabolic equations ☆

Daoqi Yang

Department of Mathematics, Wayne State University, 656 W. Kirby, # 1150, Detroit, MI 48202, USA

Received 8 March 1999; received in revised form 26 August 1999

Abstract

Dynamic finite element schemes are analyzed for second-order parabolic problems. These schemes permit different finite element spaces at different time levels in order to efficiently capture time-changing localized phenomena, such as moving sharp fronts or layers. The dynamical change of grids and interpolation polynomials is necessary and essential to many large-scale transient problems. Standard, characteristic, and mixed finite element methods with dynamic function spaces are considered for linear and nonlinear problems in a unified framework with improved a priori error estimates and convergence results. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Dynamic finite element method; Parabolic equation

1. Introduction

Many time-dependent problems involve localized phenomena, such as sharp fronts, shocks, and layers, which also change with time. The numerical simulation of these problems using the finite element method requires capabilities for efficient, dynamic, and self-adaptive local grid refinement or unrefinement and interpolation polynomial modifications.

The objective of this paper is to analyze a number of numerical schemes for parabolic problems which permit the use of different grids and different interpolation polynomials at different time levels when necessary. For many problems, such as oil reservoir and semiconductor simulation, the solution is rough in a very small region of the physical domain, but the region of roughness sweeps out a

E-mail address: yang@math.wayne.edu (D. Yang)

☆ This research was supported in part by Institute for Mathematics and its Applications at University of Minnesota.

PII: S0377-0427(99)00362-3

substantial part of the domain during the entire time period of simulation. Thus a static-grid (fixed with time) finite element method would require very fine grid over the entire domain and is often too expensive in practice. On the other hand, dynamic finite elements (changing with time) would provide great computational flexibility and efficiency, where local grid refinement and interpolation polynomial modification can be made dynamically in accordance with the changing location of singularities. With the popularity of the p and hp version finite element methods, the order of interpolation polynomials can also be adapted locally in space and dynamically in time according to the behavior of the solution. It has been proved theoretically and shown experimentally that some singularities can be resolved not by just refining local grids, but by increasing the order of approximation polynomials. A frequently encountered example is parabolic problems with nonsmooth initial data. At the beginning the solution is not smooth, fine grids and piecewise linear interpolation polynomials (i.e., the h version) may be applied. After a while the solution becomes smooth, we may use coarse grid and higher-order basis functions (i.e., the p version). For general problems, hp version finite elements may be applied to improve efficiency and accuracy; see [2,3]. If the grid and basis functions are chosen at each time level in accordance with the changing character of the solution at that time, then the dynamic finite element methods have the capability for accurately and efficiently resolving time-changing phenomena. For simplicity, however, we will consider only the h version in our analysis. The p and hp versions can be treated analogously.

Dynamic finite element schemes under the name of discontinuous Galerkin or space-time finite element methods have been discussed in [6,14–18,20,21,24,25] for model linear and nonlinear evolution problems. These methods can provide a posteriori error estimates and adaptivity based on local grid refinement at different time levels. The error analysis obtained so far is not optimal (there is a logarithmic factor of the time step size contained in the error estimates; [24,15], depends on some strong stability estimates for the discrete dual problem, and imposes some restrictions on the time and space grids. For example, the error estimates in [20,25] are nonoptimal in the sense that they contain the factor Δt^{-1}, where Δt is the time step size. The finite element spaces in [15] are required to satisfy that $S_n \subset S_{n-1}$, where S_n is the finite element space at time level $t = t_n$, or the space and time grids are required to satisfy that $h_n^2 \leqslant C\Delta t_n$, where h_n and Δt_n are the space and time grid sizes at time $t = t_n$, respectively and C is a constant. Generalization of the estimates in Eriksson and Johnson [15–18] to general variable coefficients and to nonstandard finite elements (e.g. characteristic and mixed finite elements) has not been seen.

Moving finite element method (see [4,5,29–31]) is another class of such methods which provide dynamic change of grids according to the moving local phenomena. A unique feature of moving finite elements is the inclusion of grid point movement in weak forms or in the minimization of the residual of the differential equations. That is, the position of grid points and the approximate solution at these points are solved simultaneously for each time level in such a way that the weighted residual of the differential equation, possibly with a penalty term, is minimized. This method offers a good way of solving certain kind of problems, but employ essentially the same number of grid points at all time levels and has great difficulties tackling three-dimensional problems.

A third class of dynamic finite element schemes was mainly analyzed at the theoretical level (see [13,26–28,32,38–43,46,47]), although numerical experiments were given in [43,44] based on domain decomposition and finite element discretization at each time level. Sub-optimal convergence results were derived in these papers. The idea is to follow the traditional finite differencing in time and finite element discretization in space (see [11,37]). However, since we are applying different finite element spaces at different times, the finite differencing in time is achieved by first projecting the solution from the previous time level onto the finite element space at the current time level, and then using it as initial value to compute the approximate solution at the current time level, and then using it as initial value to compute the approximate solution at the current time level. The projection is used for convergence analysis and may not need be actually computed for some of the schemes in implementation (rather, inner products must be computed for functions defined on different grids), although it is essential for the implementation of some other schemes. Computations have shown that the error propagates much more rapidly with time when the projection is replaced by interpolation, especially for wave problems [34].

In this paper, we will consider some dynamic finite element schemes which may be categorized into the third class as defined above. We will derive some improved a priori convergence estimates for general (variable and nonlinear coefficients with convection terms) parabolic problems and for general (standard, characteristic, and mixed) finite element approximations in a unified framework. This unified framework can also be applied to second-order hyperbolic problems without much modification. Although our analysis is traditional and quite simple, it applies to virtually all dynamic finite element schemes for essentially all time-dependent problems. Besides, the convergence results improve some previous ones in the literature and help provide computational insight that cannot be rigorously argued before.

We now introduce some notation that will be used throughout the paper. Let Ω denote a spatial domain in $\mathbb{R}^d$ with a piecewise uniformly Lipschitz boundary Γ. Here d is a positive integer. Denote by $H^m(\Omega)=W^{m,2}(\Omega)$ and $W^{m,p}(\Omega)$ the standard Sobolev spaces on Ω, with norms $\|\cdot\|_m$ and $\|\cdot\|_{m,p}$, respectively. Let $L^p(\Omega)$, $p=2,\infty$, denote the standard Banach spaces, with $\|\cdot\|$ denoting the L^2 norm and $\|\cdot\|_\infty$ the L^∞ norm over Ω. However, for a positive function ϕ, we use $\|\cdot\|_\phi$ to denote the weighted L^2-norm with weight function ϕ. For a normed linear space Q with norm $\|\cdot\|_Q$ and a sufficient regular function $g:[t_1,t_2]\to Q$, we define

$$\|g\|_{L^p([t_1,t_2];Q)} = \left(\int_{t_1}^{t_2} \|g(\cdot,t)\|_Q^p \,\mathrm{d}t\right)^{1/p}, \quad p=1,2,\infty,$$

with standard modification for $p=\infty$, where $[t_1,t_2]\subset[0,T]$ is a time interval, and T is a positive number. We omit $[t_1,t_2]$ from the notation when $[t_1,t_2]=[0,T]$; that is, we write $\|g\|_{L^p(Q)}$ instead of $\|g\|_{L^p([0,T];Q)}$.

We partition the time interval $[0,T]$ into $0=t_0<t_1<\cdots<t_N=T$, and denote $\Delta t_n=t_n-t_{n-1}$, $n=1,2,\dots,N$, where N is a positive integer representing the number of time steps. We will also use capital letter C, without subscripts, to denote a generic positive real constant, which may take on different values in different occurrences.

The organization of the paper is as follows. In Section 2 we give our approximation scheme and prove some convergence results for linear problems. In Section 3 we generalize the method to nonlinear problems. Then in Section 4, we consider the modified method of characteristics, and in Section 5, we treat mixed finite element methods. Finally in Section 6, we give some concluding remarks. All of our error estimates are a priori. Although a priori error estimates are not normally used to guide grid refinement and polynomial modification, they do help characterize the propagation of error due to the dynamic change of the finite element space and provide numerical schemes that have good convergence properties.

2. Linear problems

Consider the following linear parabolic problem with Dirichlet boundary condition: find $u(x,t)$ satisfying

$$\phi(x)\frac{\partial u}{\partial t} - \nabla\cdot(a(x,t)\nabla u) + b(x,t)\cdot\nabla u + c(x,t)u = f(x,t), \quad x\in\Omega,\ t\in(0,T], \tag{2.1}$$

$$u(x,t)=0, \quad x\in\Gamma,\ t\in(0,T], \tag{2.2}$$

$$u(x,0)=g(x), \quad x\in\Omega, \tag{2.3}$$

where f,g,a,b,c and ϕ are known real-valued functions, of which a is a matrix function and b is a vector function. It is assumed that ϕ and eigenvalues of a are bounded below and above by positive constants, that b and its componentwise gradient are bounded from above by positive constants, and that c is a nonnegative function. All the given functions are assumed to lie in $L^\infty(\Omega)$. In addition, assume that there exist constants μ_0 and μ_1 such that

$$0<\mu_0\leqslant\frac{1}{2}\nabla\cdot b(x,t)+c(x,t)\leqslant\mu_1, \quad (x,t)\in\Omega\times[0,T].$$

Under these conditions, problem (2.1)–(2.3) has a unique solution.

Our numerical method will allow us to apply different finite element spaces at different times in order to capture moving local phenomena. For $n=0,1,2,\ldots,N$, let $T_n=\{K\}$ be a spatial discretization of the domain Ω, and S_n be a finite element subspace of $H_0^1(\Omega)$ with grid parameter $h_{n,K}$, and interpolation polynomials of degree $k_{n,K}$ in element K and at time step n. Also let

$$h_n=\max_{K\in T_n}\{h_{n,k}\}, \qquad k_n=\max_{K\in T_n}\{k_{n,K}\} \quad\text{and}\quad k=\max_n\{k_n\}.$$

We assume that the following approximation property holds: for $n=0,1,2,\ldots,N$,

$$\inf_{v\in S_n}\|w-v\|_j\leqslant C\left(\sum_{K\in T_n}h_{n,K}^{2(k_{n,K}+1-j)}\|w\|^2_{H^{k_{n,K}+1}(K)}\right)^{1/2}, \quad j=0,1,\ \forall w\in H_0^1(\Omega)\cap H^{k_n+1}(\Omega), \tag{2.4}$$

where C is a constant independent of w,n,h_n, and k. Note that in the case of p and hp versions, the constant C depends on k_n. Here we assume that piecewise polynomials of degree less than a certain number k are used, while still having the flexibility of changing the polynomial degrees slightly from time level to time level, when necessary.

We first define the implicit Euler scheme. Suppose that $U_0\in S_0$ is an initial approximation of $u(\cdot,0)$, we define our first dynamic finite element scheme as follows:

Algorithm 2.1. *For $n=1,2,\ldots,N$, first compute the weighted L^2 projection $\hat U_{n-1}\in S_n$ by solving*

$$(\phi(\hat U_{n-1}-U_{n-1}),v)=0, \quad \forall v\in S_n; \tag{2.5}$$

then compute $U_n\in S_n$ by

$$\left(\phi\frac{U_n-\hat U_{n-1}}{\Delta t_n},v\right)+(a_n\nabla U_n,\nabla v)+(b_n\cdot\nabla U_n+c_nU_n,v)=(f_n,v), \quad \forall v\in S_n, \tag{2.6}$$

where $(f,g)=\int_\Omega f\cdot g\,\mathrm{d}x$ and $\xi_n=\xi(x,t_n)$ for any function ξ.

Below are some remarks about scheme (2.5)–(2.6). Eq. (2.5) gives a weighted L^2 projection $\hat{U}_{n-1}$ of the previous approximate solution U_{n-1} onto the current finite element space S_n when different finite element spaces are used at times $t = t_n$ and t_{n-1}. This projection is used in (2.6) as initial value to calculate U_n, the approximate solution at $t = t_n$. When the finite element space remains unchanged for all time levels, scheme (2.5)–(2.6) reduces to the standard one [11,37]. Note that Algorithm 2.1 is very similar to the space–time finite element scheme in Eriksson and Johnson [15] with piecewise constant polynomials in time (whose a priori error estimates can be improved by the argument presented in this paper), and is the same as a scheme considered in [13,26].

The Crank–Nicolson scheme can be defined in the standard way.

Algorithm 2.2. *For $n = 1, 2, \ldots, N$, compute first $\hat{U}_{n-1} \in S_n$ by (2.5) and then $U_n \in S_n$ by*

$$\left(\phi\frac{U_n - \hat{U}_{n-1}}{\Delta t_n}, v\right) + \left(a_{n-1/2}\nabla\frac{U_n + \hat{U}_{n-1}}{2}, \nabla v\right) + \left(b_{n-1/2}\cdot\nabla\frac{U_n + \hat{U}_{n-1}}{2} + c_{n-1/2}\frac{U_n + \hat{U}_{n-1}}{2}, v\right) = (f_{n-1/2}, v), \quad \forall v \in S_n, \tag{2.7}$$

where $\xi_{n-1/2} = \xi(x, t_{n-1/2})$ for any function ξ, and $t_{n-1/2} = \frac{1}{2}(t_n + t_{n-1})$.

It should be noted that the projection may not have to be explicitly computed in Algorithm 2.1, but rather an inner product has to be calculated for functions defined on new and old grids. However, this projection is necessary and must be explicitly computed in Algorithm 2.2 since it enters the H^1 inner product.

In order to make error estimates for the schemes, we make use of the elliptic projection $R_n u$ of u: find $R_n u(x,t) \in S_n$ for each $t \in [0, T]$ such that

$$(a_n\nabla(u - R_n u)(\cdot, t), \nabla u) + (c_n(u - R_n u)(\cdot, t), v) = 0, \quad \forall v \in S_n. \tag{2.8}$$

Similar to (2.4), we assume that

$$\|u - R_n u\|_j \leqslant C\left(\sum_{K\in T_n} h_{n,K}^{2(k_{n,K}+1-j)}\|u\|^2_{H^{k_{n,K}+1}(K)}\right)^{1/2}, \quad j = 0, 1, \ \forall t \in [0, T]. \tag{2.9}$$

This can be guaranteed by requiring that the triangulation be regular at each time level and that all finite elements be affine. A proof can be found in [9,7] with a slight modification that duality argument need be applied on each element to obtain the optimal L^2-norm estimate for nonuniform grids.

We now state and prove the following convergence estimates for the implicit Euler scheme.

Theorem 2.1. *Suppose that the solution u to problem (2.1)–(2.3) satisfies the regularity requirement: $u, \partial u/\partial t \in L^\infty(H^{k+1}(\Omega) \cap H_0^1(\Omega))$, $\partial^2 u/\partial t^2 \in L^1(L^2(\Omega))$. Let U_n be the solution of scheme*

(2.5)–(2.6). *Then we have the error estimates for* $m = 1, 2, \ldots, N$, *when* Δt_n *are sufficiently small,*

$$\max_{1\leqslant n\leqslant m} \|u_n - U_n\|^2 \leqslant C\left\{E_n + \max_{0\leqslant n\leqslant m} \sum_{K\in T_n} h_{n,K}^{2(k_{n,K}+1)} \|u(\cdot,t_n)\|^2_{H^{k_{n,K}+1}(K)}\right\}, \tag{2.10}$$

$$\sum_{n=1}^{m} \Delta t_n[(a_n\nabla(u_n - U_n), \nabla(u_n - U_n)) + (c_n(u_n - U_n), u_n - U_n)]$$

$$\leqslant C\left\{E_n + \sum_{n=1}^{m}\sum_{K\in T_n} \Delta t_n h_{n,K}^{2k_{n,K}} \|u(\cdot,t_n)\|^2_{H^{k_{n,K}+1}(K)}\right\}, \tag{2.11}$$

where

$$E_n = \|u_0 - U_0\|^2 + \left[\sum_{n=1}^{m}\left(\sum_{K\in T_n}\left[\int_{t_{n-1}}^{t_n} h_{n,K}^{k_{n,K}+1}\left\|\frac{\partial u}{\partial t}\right\|_{H^{k_{n,K}+1}(K)} \mathrm{d}t\right]^2\right)^{1/2}\right]^2$$

$$+\sum_{n=1}^{m}\Delta t_n \sum_{K\in T_n} h_{n,K}^{2(k_{n,K}+1)}\|u(\cdot,t_n)\|^2_{H^{k_{n,K}+1}(K)}$$

$$+\sum_{n=1}^{m}\left[\Delta t_n\int_{t_{n-1}}^{t_n}\left\|\frac{\partial^2 u}{\partial t^2}\right\|\mathrm{d}t\right]^2 + \left[\sum_{n=1}^{m}\|(R_n - R_{n-1})u_{n-1}\|\right]^2, \tag{2.12}$$

and R_n *is the elliptic projection defined by* (2.8).

Proof. We will use the following notation in the proof:

$$e_n = U_n - R_n u_n, \quad \hat{e}_{n-1} = \hat{U}_{n-1} - R_n u_{n-1},$$

$$r_n = u_n - R_n u_n, \quad \hat{r}_{n-1} = u_{n-1} - R_n u_{n-1}.$$

Notice that the exact solution u satisfies

$$\left(\phi\frac{u_n - u_{n-1}}{\Delta t_n}, v\right) + (a_n\nabla u_n, \nabla v) + (b_n\cdot\nabla u_n + c_n u_n, v)$$

$$= (f_n, v) + (\phi\rho_n, v), \quad \forall v\in H_0^1(\Omega), \tag{2.13}$$

where

$$\|\rho_n\| = \left\|\frac{u_n - u_{n-1}}{\Delta t_n} - \frac{\partial u}{\partial t}(t_n)\right\| \leqslant \int_{t_{n-1}}^{t_n}\left\|\frac{\partial^2 u}{\partial t^2}\right\|\mathrm{d}t. \tag{2.14}$$

Subtracting (2.13) from (2.6) and using (2.8) yield

$$\left(\phi\frac{e_n - \hat{e}_{n-1}}{\Delta t_n}, v\right) + (a_n\nabla e_n, \nabla v) + (c_n e_n, v)$$

$$= \left(\phi\left(\frac{r_n - \hat{r}_{n-1}}{\Delta t_n} - \rho_n\right), v\right) + (b_n\cdot\nabla(u_n - U_n), v), \quad \forall v\in S_n. \tag{2.15}$$

Letting $v = e_n$ in (2.15) we obtain the error equation

$$\left(\phi\frac{e_n - \hat{e}_{n-1}}{\Delta t_n}, e_n\right) + (a_n \nabla e_n, \nabla e_n) + (c_n e_n, e_n)$$

$$= \left(\phi\left(\frac{r_n - \hat{r}_{n-1}}{\Delta t_n} - \rho_n\right), e_n\right) + (b_n \cdot \nabla(u_n - U_n), e_n). \tag{2.16}$$

It is easy to see that

$$\left(\phi\frac{e_n - \hat{e}_{n-1}}{\Delta t_n}, e_n\right) = \frac{\|e_n\|_\phi^2 - \|\hat{e}_{n-1}\|_\phi^2}{2\Delta t_n} + \frac{1}{2\Delta t_n}\|e_n - \hat{e}_{n-1}\|_\phi^2 \tag{2.17}$$

and

$$\|r_n - \hat{r}_{n-1}\| = \|(I - R_n)(u_n - u_{n-1})\|$$

$$\leqslant C\left(\sum_{K\in T_n} h_{n,K}^{2(k_{n,K}+1)}\|u_n - u_{n-1}\|_{H^{k_{n,K}+1}(K)}^2\right)^{1/2}$$

$$= C\left(\sum_{K\in T_n} h_{n,K}^{2(k_{n,K}+1)}\left\|\int_{t_{n-1}}^{t_n}\frac{\partial u}{\partial t}\,\mathrm{d}t\right\|_{H^{k_{n,K}+1}(K)}^2\right)^{1/2}$$

$$\leqslant C\left(\sum_{K\in T_n}\left[\int_{t_{n-1}}^{t_n} h_{n,K}^{k_{n,K}+1}\left\|\frac{\partial u}{\partial t}\right\|_{H^{k_{n,K}+1}(K)}\mathrm{d}t\right]^2\right)^{1/2}. \tag{2.18}$$

Applying integration by parts and the ε-inequality, we have

$$(b_n \cdot \nabla(u_n - U_n), e_n) = (b_n \cdot \nabla r_n, e_n) - (b_n \cdot \nabla e_n, e_n)$$

$$= -(\nabla \cdot (e_n b_n), r_n) + \int_\Gamma e_n r_n b_n \cdot v\,\mathrm{d}s - (b_n \cdot \nabla e_n, e_n)$$

$$= -(\nabla \cdot (e_n b_n), r_n) - (b_n \cdot \nabla e_n, e_n)$$

$$\leqslant \tfrac{1}{2}(a_n \nabla e_n, \nabla e_n) + C(\|e_n\|_\phi^2 + \|r_n\|^2), \tag{2.19}$$

where v is the unit outward normal of Γ.

Combining (2.16)–(2.19), and (2.14), we have the following error inequality:

$$\|e_n\|_\phi^2 - \|\hat{e}_{n-1}\|_\phi^2 + \Delta t_n[(a_n \nabla e_n, \nabla e_n) + (c_n e_n, e_n)] \leqslant C[F_n\|e_n\|_\phi + \Delta t_n(\|e_n\|_\phi^2 + \|r_n\|^2)], \tag{2.20}$$

where

$$F_n = \left(\sum_{K\in T_n}\left[\int_{t_{n-1}}^{t_n} h_{n,K}^{k_{n,K}+1}\left\|\frac{\partial u}{\partial t}\right\|_{H^{k_{n,K}+1}(K)}\mathrm{d}t\right]^2\right)^{1/2} + \Delta t_n\int_{t_{n-1}}^{t_n}\left\|\frac{\partial^2 u}{\partial t^2}\right\|\mathrm{d}t. \tag{2.21}$$

We now find the relationship between $\|e_n\|_\phi$ and $\|\hat{e}_n\|_\phi$. Eq. (2.5) implies that

$$(\phi(\hat{e}_{n-1} - e_{n-1}), v) = (\phi(\hat{r}_{n-1} - r_{n-1}), v), \quad \forall v \in S_n.$$

Choosing $v = \hat{e}_{n-1}$ and using Schwarz's inequality, we have

$$\|\hat{e}_{n-1}\|_\phi^2 - \|e_{n-1}\|_\phi^2 \leqslant 2\|\hat{e}_{n-1}\|_\phi \|\hat{r}_{n-1} - r_{n-1}\|_\phi \leqslant \tfrac{1}{2}\|\hat{e}_{n-1}\|_\phi^2 + 2\|\hat{r}_{n-1} - r_{n-1}\|_\phi^2,$$

from which we derive that

$$\|\hat{e}_{n-1}\|_\phi^2 \leqslant 2\|e_{n-1}\|_\phi^2 + 4\|\hat{r}_{n-1} - r_{n-1}\|_\phi^2 \tag{2.22}$$

and

$$\|\hat{e}_{n-1}\|_\phi^2 - \|e_{n-1}\|_\phi^2 \leqslant 2\sqrt{2}\|e_{n-1}\|_\phi \|\hat{r}_{n-1} - r_{n-1}\|_\phi + 4\|\hat{r}_{n-1} - r_{n-1}\|_\phi^2. \tag{2.23}$$

Combining (2.20) and (2.23) we see that

$$\begin{aligned} &\|e_n\|_\phi^2 - \|e_{n-1}\|_\phi^2 + \Delta t_n[(a_n \nabla e_n, \nabla e_n) + (c_n e_n, e_n)] \\ &\quad \leqslant C[F_n\|e_n\|_\phi + \Delta t_n(\|e_n\|_\phi^2 + \|r_n\|^2) + \|r_{n-1} - \hat{r}_{n-1}\|_\phi \|e_{n-1}\|_\phi + \|r_{n-1} - \hat{r}_{n-1}\|_\phi^2]. \end{aligned} \tag{2.24}$$

Summing (2.24) from $n = 1$ to m $(1 \leqslant m \leqslant N)$, we obtain

$$\begin{aligned} &\max_{1\leqslant n\leqslant m} \|e_n\|_\phi^2 - \|e_0\|_\phi^2 + \sum_{n=1}^m \Delta t_n[(a_n \nabla e_n, \nabla e_n) + (c_n e_n, e_n)] \\ &\quad \leqslant C\sum_{n=1}^m (F_n\|e_n\|_\phi + \|r_{n-1} - \hat{r}_{n-1}\|_\phi \|e_{n-1}\|_\phi + \|r_{n-1} - \hat{r}_{n-1}\|_\phi^2 + \Delta t_n(\|e_n\|_\phi^2 + \|r_n\|^2)) \\ &\quad \leqslant C\left\{\max_{0\leqslant n\leqslant m}\|e_n\|_\phi \sum_{n=1}^m (F_n + \|r_{n-1} - \hat{r}_{n-1}\|_\phi) + \sum_{n=1}^m [\|r_{n-1} - \hat{r}_{n-1}\|_\phi^2 + \Delta t_n(\|e_n\|_\phi^2 + \|r_n\|^2)]\right\} \\ &\quad \leqslant \frac{1}{2}\max_{0\leqslant n\leqslant m}\|e_n\|_\phi^2 + C\left\{\left[\sum_{n=1}^m F_n\right]^2 + \left[\sum_{n=1}^m \|r_{n-1} - \hat{r}_{n-1}\|_\phi\right]^2 + \sum_{n=1}^m \Delta t_n(\|e_n\|_\phi^2 + \|r_n\|^2)\right\}. \end{aligned}$$

Then

$$\begin{aligned} &\frac{1}{2}\max_{0\leqslant n\leqslant m}\|e_n\|_\phi^2 - \frac{3}{2}\|e_0\|_\phi^2 + \sum_{n=1}^m \Delta t_n[(a_n \nabla e_n, \nabla e_n) + (c_n e_n, e_n)] \\ &\quad \leqslant C\sum_{n=1}^m \Delta t_n\|e_n\|_\phi^2 + C\left\{\left[\sum_{n=1}^m F_n\right]^2 + \left[\sum_{n=1}^m \|r_{n-1} - \hat{r}_{n-1}\|_\phi\right]^2 + \sum_{n=1}^m \Delta t_n\|r_n\|^2\right\}. \end{aligned}$$

Applying the discrete Gronwall Lemma, triangular inequality, and (2.9) completes the proof of the theorem. □

Corollary 2.1. *When a static finite element space is used for $0 \leqslant t \leqslant T$, i.e., when $h_n = h$, $k_n = k$ for $n = 0, 1, \ldots, N$, we have the error estimates*

$$\|u_m - U_m\| = \mathrm{O}(\Delta t + h^{k+1} + \|u_0 - U_0\|). \tag{2.25}$$

Thus the implicit Euler algorithm analyzed in [37] *is the $h_n = h$ and $k_n = k$ case of scheme* (2.5)–(2.6).

The results in Theorem 2.1 show that our error estimation consists of three parts: an optimal temporal finite difference discretization error, an optimal spatial finite element discretization error, and an error term due to the projections of the approximated solution from old finite element spaces onto new finite element spaces. The error term due to projection in the worst case can be bounded by

$$\begin{aligned}\left[\sum_{n=1}^{m}\|(R_n-R_{n-1})u_{n-1}\|\right]^2 &= \left[\sum_{n=1}^{m}\|r_{n-1}-\hat{r}_{n-1}\|\right]^2 \leqslant \left[\sum_{n=1}^{m}\delta_{n-1}\{\|r_{n-1}\|+\|\hat{r}_{n-1}\|\}\right]^2\\ &\leqslant \left[\sum_{n=1}^{m}\delta_{n-1}\left\{\left(\sum_{K\in T_{n-1}} h_{n-1,K}^{2(k_{n-1,K}+1)}\|u(\cdot,t_{n-1})\|^2_{H^{k_{n-1,K}+1}(K)}\right)^{1/2}\right.\right.\\ &\qquad\left.\left.+\left(\sum_{K\in T_n} h_{n,K}^{2(k_{n,K}+1)}\|u(\cdot,t_{n-1})\|^2_{H^{k_{n,K}+1}(K)}\right)^{1/2}\right\}\right]^2,\end{aligned} \tag{2.26}$$

where $\delta_n=0$ if $S_n=S_{n+1}$ and $\delta_n=1$ otherwise.

Thus every change of the finite element space contributes to the error estimates by an amount that is not explicitly dependent upon the time step; but this amount depends on the way that the finite element space changes. The closer is the space S_n to S_{n-1}, the smaller is the error amount of contribution. This is intuitive but cannot be observed from some previous error estimates. Theorem 2.1 is also related to early results [26–28,32,13,15,16,38–42,44,46,47]. But our results do not include M, the number of different finite element spaces applied up to the current time level, or some logarithmic factor of the inverse of the time step size. To be more specific, let us state the error estimate in Theorem 8.2 of Johnson [24]:

$$\|u_n-U_n\|\leqslant C\left(1+\log\frac{t_n}{\Delta t_n}\right)^{1/2}\left(\max_{m\leqslant n}\int_{t_{m-1}}^{t_m}\left\|\frac{\mathrm{d}u(s)}{\mathrm{d}t}\right\|\,\mathrm{d}s+\max_{t\leqslant t_n}h^2\|u(t)\|_{H^2(\Omega)}\right), \tag{2.27}$$

under the assumption that $\Delta t_{n-1}\leqslant\gamma\Delta t_n$, and that piecewise linear polynomials in space are applied, where C and γ are constants. Our estimates show that the error due to the dynamic change of the finite element space should be independent of the time step size Δt_n. Thus the logarithmic growth factor in (2.27) should not be there in general and makes estimate (2.27) nonoptimal. In fact, estimate (2.27) implies that the factor $\log(|t_n/\Delta t_n|)$ is there even if the finite element space changes only once throughout the entire time interval $[0,T]$, which clearly indicates a nonsharp estimate. Note also that the error due to time discretization in (2.27) is bounded by

$$\left(1+\log\frac{t_n}{\Delta t_n}\right)^{1/2}\max_{m\leqslant n}\int_{t_{m-1}}^{t_m}\left\|\frac{\mathrm{d}u(s)}{\mathrm{d}t}\right\|\,\mathrm{d}s.$$

The results of our Theorem 2.1 have optimal time discretization errors, but requiring the existence of the second-order derivative in time, while (2.27) requires only first-order derivative in time. Thus both our results here and previous ones [24] have advantages and disadvantages. In particular, results of Theorem 8.2 in [24] have not seen being extended to general variable coefficients and mixed and characteristic finite element methods, while our results can be easily generalized to these situations without restrictive assumptions such as $\Delta t_{n-1} \leqslant \gamma \Delta t_n$.

Our results are also different from those obtained by Dupont [13] and Bank and Santos [6], in which our error estimates are given in standard norms independent of the finite element grids. Also, the finite element grids in our method are not required to change continuously in any fashion. Note that the error estimates in [13,6] involve grid-dependent norms in the $H^{-1}(\Omega)$ sense, which are hard to compute. Similar results in [13] were obtained in [22]. A one-dimensional problem was analyzed in [23].

The Crank–Nicolson scheme (Algorithm 2.2) can be analyzed following the steps in the proof of Theorem 2.1. We omit the analysis here and just present the following theorem without proof:

Theorem 2.2. *Suppose that the solution u to problem* (3.1)–(3.3) *satisfies the regularity condition*: $u, \partial u/\partial t \in L^\infty(H^{k+1}(\Omega) \cap H_0^1(\Omega)), \partial^2 u/\partial t^2 \in L^1(L^2(\Omega))$. *Let* U_n *be the solution of Crank–Nicolson scheme* (2.7). *Then we have the error estimates for* $m = 1, 2, \ldots, N$,

$$
\begin{aligned}
&\max_{1 \leqslant n \leqslant m} \|u_n - U_n\|^2 \\
&\leqslant C \left\{ \|u_0 - U_0\|^2 + \left[\sum_{n=1}^{m} \left(\sum_{K \in T_n} \left[\int_{t_{n-1}}^{t_n} h_{n,K}^{k_{n,K}+1} \left\| \frac{\partial u}{\partial t} \right\|_{H^{k_{n,K}+1}(K)} \mathrm{d}t \right]^2 \right)^{1/2} \right]^2 \right. \\
&\quad + \sum_{n=1}^{m} \Delta t_n \left(\sum_{K \in T_{n-1}} h_{n-1,K}^{2(k_{n-1,K}+1)} \|u(\cdot, t_{n-1})\|^2_{H^{k_{n-1,K}+1}(K)} + \sum_{K \in T_n} h_{n,K}^{2(k_{n,K}+1)} \|u(\cdot, t_n)\|^2_{H^{k_{n,K}+1}(K)} \right) \\
&\quad + \max_{0 \leqslant n \leqslant m} \sum_{K \in T_n} h_{n,K}^{2(k_{n,K}+1)} \|u(\cdot, t_n)\|^2_{H^{k_{n,K}+1}(K)} + \left[\sum_{n=1}^{m} \Delta t_n^2 \left\| \frac{\partial^3 u}{\partial t^3} \right\|_{L^1([t_{n-1}, t_n]; L^2(\Omega))} \right]^2 \\
&\quad + \sum_{n=1}^{m} \Delta t_n^5 \left(\|u\|^2_{L^\infty(H_0^1(\Omega))} + \left\| \frac{\partial u}{\partial t} \right\|^2_{L^\infty(H_0^1(\Omega))} + \left\| \frac{\partial^2 u}{\partial t^2} \right\|^2_{L^\infty(H_0^1(\Omega))} \right) \\
&\quad \left. + \left[\sum_{n=1}^{m} \|(R_n - R_{n-1}) u_{n-1}\| \right]^2 \right\}. \qquad (2.28)
\end{aligned}
$$

Compared to the implicit Euler scheme, this scheme has second-order accuracy in time, but requires higher-order derivatives in time. Dupont [13, p. 92] claimed that an analysis of the Crank–Nicolson scheme was not possible in his framework. Thus our analysis here, although simple, not only gives improved error estimates, but also provides software implementors with computational insights and more theoretically guaranteed convergent numerical schemes, on which some previous theory had remained silent.

3. Nonlinear problems

Consider the following quasilinear parabolic problem with Dirichlet boundary condition: find $u(x,t)$ satisfying

$$\phi(x)\frac{\partial u}{\partial t} - \nabla\cdot(a(x,u)\nabla u) + b(x,u)\cdot\nabla u = f(x,u), \quad x\in\Omega, \ t\in(0,T], \tag{3.1}$$

$$u(x,t) = 0, \quad x\in\Gamma, \ t\in(0,T], \tag{3.2}$$

$$(x,0) = g(x), \quad x\in\Omega, \tag{3.3}$$

whose weak formulation can be put into the fashion: find a differentiable mapping $u:[0,T]\to H_0^1(\Omega)$ such that

$$(\phi u_t, v) + A(u;u,v) + B(u;u,v) = (f(u),v), \quad \forall v\in H_0^1(\Omega), \tag{3.4}$$

$$(u(\cdot,0),v) = (g,v), \quad \forall v\in H_0^1(\Omega), \tag{3.5}$$

where $(\cdot,\cdot)$ denotes the inner product in $L^2(\Omega), f(u)=f(x,u)$, and

$$A(w;u,v) = (a(x,w)\nabla u, \nabla v), \quad B(w;u,v) = (b(x,w)\cdot\nabla u, v). \tag{3.6}$$

Let S_n be a finite element space at time $t=t_n$ satisfying (2.4). We make some assumptions on the coefficients:

$$\text{For } (x,v)\in\Omega\times\mathscr{R}, \quad C_1\|\eta\|^2 \leqslant \sum_{i,j=1}^{d}(a_{ij}(x,v)\eta_i,\eta_j) \leqslant C_2\|\eta\|^2, \quad \max_{1\leqslant i\leqslant d}|b_i(x,v)| \leqslant C_2, \tag{3.7}$$

$$f, b_i, a_{ij}, \text{ are uniformly Lipschitz continuous with respect to their } (d+1)\text{th variable}, \tag{3.8}$$

$$\text{For } 1\leqslant i\leqslant d, \frac{\partial b_i(x,v)}{\partial x_i} \text{ exists and is bounded from above}, \tag{3.9}$$

$$f(\cdot,0)\in L^2(\Omega), \quad g\in H^{k+1}(\Omega)\cap H_0^1(\Omega), \tag{3.10}$$

$$u \text{ is unique to } (3.1)-(3.3), \text{ and } u, \frac{\partial u}{\partial t} \in L^\infty(H^{k+1}(\Omega)\cap H_0^1(\Omega)), \frac{\partial^2 u}{\partial t^2}\in L^1(L^2(\Omega)). \tag{3.11}$$

Suppose that $U_0\in S_0$ is an initial approximation of $u(\cdot,0)$, we define our dynamic finite element scheme with a parameter θ as follows:

Algorithm 3.1. *For $n=1,2,\dots,N$, first compute the weighted L^2 projection $\hat U_{n-1}\in S_n$ by solving*

$$(\phi(\hat U_{n-1} - U_{n-1}), v) = 0, \quad \forall v\in S_n; \tag{3.12}$$

then compute $U_n\in S_n$ by

$$\left(\phi\frac{U_n - \hat U_{n-1}}{\Delta t_n}, v\right) + A(\hat U_{n,\theta}; \hat U_{n,\theta}, v) + B(\hat U_{n,\theta}; \hat U_{n,\theta}, v) = (f(\hat U_{n,\theta}), v), \quad \forall v\in S_n, \tag{3.13}$$

where $\hat U_{n,\theta} = \frac{1}{2}(1+\theta)U_n + \frac{1}{2}(1-\theta)\hat U_{n-1}, \quad 0\leqslant\theta\leqslant 1.$

Using Brouwer's fixed-point theorem, we can show that scheme (3.12)–(3.13) has a solution for sufficiently small Δt_n (see [11]). Note that $\theta=0$ corresponds to the Crank–Nicolson scheme, while $\theta=1$ corresponds to the implicit Euler scheme. We have the following error estimate for the θ-scheme:

Theorem 3.1. *Suppose that the solution u to problem* (3.1)–(3.3) *satisfies* (3.11). *Let U_n be the solution of scheme* (3.12)–(3.13). *Then we have the error estimates for $m=1,2,\dots,N$,*

$$\begin{aligned}
&\max_{1\leqslant n\leqslant m}\|u_n-U_n\|^2\\
&\leqslant C\Bigg\{\|u_0-U_0\|^2+\left[\sum_{n=1}^{m}\left(\sum_{K\in T_n}\left[\int_{t_{n-1}}^{t_n}h_{n,K}^{k_{n,K}+1}\left\|\frac{\partial u}{\partial t}\right\|_{H^{k_{n,K}+1}(K)}\mathrm{d}t\right]^2\right)^{1/2}\right]^2\\
&\quad+\sum_{n=1}^{m}\Delta t_n\left(\sum_{K\in T_{n-1}}h_{n-1,K}^{2(k_{n-1,K}+1)}\|u(\cdot,t_{n-1})\|^2_{H^{k_{n-1,K}+1}(K)}+\sum_{K\in T_n}h_{n,K}^{2(k_{n-1,K}+1)}\|u(\cdot,t_n)\|^2_{H^{k_{n,K}+1}(K)}\right)\\
&\quad+\max_{0\leqslant n\leqslant m}\sum_{K\in T_n}h_{n,K}^{2(k_{n,K}+1)}\|u(\cdot,t_n)\|^2_{H^{k_{n,K}+1}(K)}+\sum_{n=1}^{m}\Delta t_n^3\left(\|u\|^2_{L^\infty(H_0^1(\Omega))}+\left\|\frac{\partial u}{\partial t}\right\|^2_{L^\infty(H_0^1(\Omega))}\right)\\
&\quad+\left[\sum_{n=1}^{m}\Delta t_n\left\|\frac{\partial^2 u}{\partial t^2}\right\|_{L^1([t_{n-1},t_n];L^2(\Omega))}\right]^2+\left[\sum_{n=1}^{m}\|(R_n-R_{n-1})u_{n-1}\|\right]^2\Bigg\}.
\end{aligned}\tag{3.14}$$

The proof of this theorem is omitted but can be easily done along the lines presented in the previous section. It states that this θ-algorithm (3.12)–(3.13) has first-order accuracy in time. When $\theta=0$, a second-order accuracy in time can be proved by enhancing some estimates in the proof of Theorem 3.1.

Note scheme (3.12)–(3.13) is nonlinear and requires some linearization technique such as Newton's iteration. However, a first-order linear scheme can be defined in a standard way.

Algorithm 3.2. *For $n=1,2,\dots,N$, compute first $\hat U_{n-1}\in S_n$ by* (3.12) *and then $U_n\in S_n$ by*

$$\left(\phi\frac{U_n-\hat U_{n-1}}{\Delta t_n},v\right)+A(\hat U_n;\hat U_{n,\theta},v)+B(\hat U_n;\hat U_{n,\theta},v)=(f(\hat U_n),v),\quad \forall v\in S_n. \tag{3.15}$$

A class of predictor–corrector schemes, which are second order in time for $\theta=0$ and first order otherwise, can be defined as follows.

Algorithm 3.3. *For $n=1,2,\dots,N$, compute first $\hat U_{n-1}\in S_n$ by* (3.12) *and then $W_n\in S_n$ and $U_n\in S_n$*

by

$$\left(\phi\frac{W_n-\hat{U}_{n-1}}{\Delta t_n},v\right)+A(\hat{U}_n;\hat{W}_{n,\theta},v)+B(\hat{U}_n;\hat{W}_{n,\theta},v)=(f(\hat{U}_n),v),\quad \forall v\in S_n, \tag{3.16}$$

$$\left(\phi\frac{U_n-\hat{U}_{n-1}}{\Delta t_n},v\right)+A(\hat{W}_{n,\theta};\hat{U}_{n,\theta},v)+B(\hat{W}_{n,\theta};\hat{U}_{n,\theta},v)=(f(\hat{W}_{n,\theta}),v),\quad \forall v\in S_n, \tag{3.17}$$

where $\hat{W}_{n,\theta}=\frac{1}{2}(1+\theta)W_n+\frac{1}{2}(1-\theta)\hat{U}_{n-1}$.

An extrapolated scheme can be defined by the following algorithm.

Algorithm 3.4. *Given two initial approximations* U_0 *and* U_1, *for* $n=2,3,\ldots,N$, *compute first* $\hat{U}_{n-1}\in S_n$ *by* (3.12) *and then* $U_n\in S_n$ *by*

$$\left(\phi\frac{U_n-\hat{U}_{n-1}}{\Delta t_n},v\right)+A(\tilde{U}_n;\hat{U}_{n,\theta},v)+B(\tilde{U}_n;\hat{U}_{n,\theta},v)=(f(\tilde{U}_{n-1}),v),\quad \forall v\in S_n, \tag{3.18}$$

where

$$\tilde{U}_{n-1}=\left(1+\frac{\Delta t_n}{2\Delta t_{n-1}}\right)U_{n-1}-\frac{\Delta t_n}{2\Delta t_{n-1}}U_{n-2}.$$

Compared to the second-order-in-time space-time finite element schemes in [15], Algorithms 3.1, 3.3 and 3.4 do not lead to a coupled system like (1.1) in [15] or (8.38) in [24]. We have chosen Crank–Nicolson scheme as an example to show that our simple technique applies to virtually all time-dependent problems without much difficulty adapting from one problem to another, although Crank–Nicolson scheme may not be the best choice for all transient problems. Note the technique in [15] is not easily generalized to other problems [16–18,25], even for linear general coefficients.

4. A dynamic characteristic finite element scheme

For convection-dominated diffusion problems, the modified method of characteristics may be preferred; see [12,19,33,6,40,45]. In this method, time discretization is made along or near the characteristic direction, instead of the t direction for standard finite difference methods. In this section, we will see that our technique applies almost trivially to characteristic finite element methods. It seems that the technique in [15–18,24,25] has not been applied to characteristic finite element methods. Bank and Santos generalized Dupont [13] to this case with mesh-dependent norms and certain restrictive assumptions on the mesh change. The results in this section also improve early ones by the author [40].

Consider the linear problem (2.1)–(2.3) as an example. Define the characteristic direction $\tau(x)$ as

$$\frac{\partial}{\partial\tau}=\frac{1}{\sqrt{\phi(x)^2+|b(x,t)|^2}}\left(\phi\frac{\partial}{\partial t}+b(x,t)\cdot\nabla\right). \tag{4.1}$$

Thus, Eq. (2.1) can be rewritten in the form

$$\sqrt{\phi^2+|b|^2}\,\frac{\partial u}{\partial\tau}-\nabla\cdot(a\nabla u)+cu=f. \tag{4.2}$$

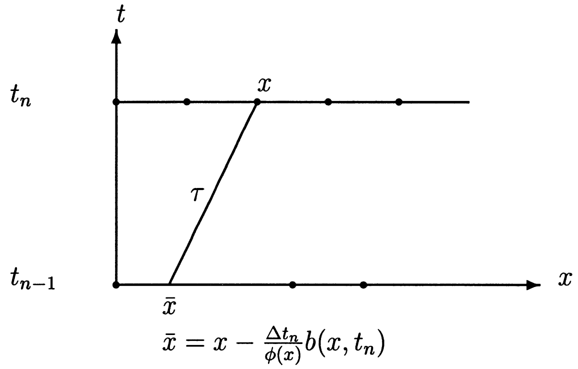

Fig. 1. Discretization along the characteristic line, where x is a grid point while $\bar{x}$ is not.

Since

$$\bar{x} = x - \frac{b(x,t_n)}{\phi(x)}\Delta t_n \tag{4.3}$$

approximates the characteristic through (x,t_n) by its tangent at (x,t_n), as in Fig. 1, we have the following backward-difference approximation:

$$\begin{aligned}
&\sqrt{\phi^2(x) + |b(x,t_n)|^2}\,\frac{\partial u}{\partial \tau}(x,t_n) \\
&\approx \sqrt{\phi^2(x) + |b(x,t_n)|^2}\,\frac{u(x,t_n) - u(\bar{x},t_{n-1})}{[|x-\bar{x}|^2 + \Delta t_n^2]^{1/2}} \\
&= \phi(x)\frac{u(x,t_n) - u(\bar{x},t_{n-1})}{\Delta t_n}.
\end{aligned} \tag{4.4}$$

Then the implicit Euler scheme along characteristics reads:

Algorithm 4.1. *For $n = 1,2,\ldots,N$ compute first $\hat{U}_{n-1} \in S_n$ by* (2.5) *and then $U_n \in S_n$ by*

$$\left(\phi\frac{U_n - \hat{\bar{U}}_{n-1}}{\Delta t_n}, v\right) + (a_n \nabla U_n, \nabla v) + (c_n U_n, v) = (f_n, v), \quad \forall v \in S_n, \tag{4.5}$$

where $\bar{U}_{n-1} = U(\bar{x}, t_{n-1})$, $\hat{\bar{U}}_{n-1} = \hat{U}(\bar{x}, t_{n-1})$, $\bar{x}$ is defined by (4.3). *Near the boundary of the domain, the characteristic curve may trace out of the domain. Thus periodicity of the solution function is required or the velocity vector b is assumed to vanish near the boundary. Otherwise some reflection principle need be applied.*

Theorem 4.1. *Suppose that the solution u to problem* (2.1)–(2.3) *satisfies the regularity condition: u, $\partial u/\partial t \in L^\infty(H^{k+1}(\Omega) \cap H_0^1(\Omega)), \partial^2 u/\partial \tau^2 \in L^1(L^2(\Omega))$. Let U_n be the solution of Algorithm* 4.1.

Then we have the error estimates for $m=1,2,\dots,N$,

$$\max_{1\leqslant n\leqslant m}\|u_n-U_n\|^2$$
$$\leqslant C\left\{\|u_0-U_0\|^2+\left[\sum_{n=1}^{m}\left(\sum_{K\in T_n}\left[\int_{t_{n-1}}^{t_n} h_{n,K}^{k_{n,K}+1}\left\|\frac{\partial u}{\partial t}\right\|_{H^{k_{n,K}+1}(K)}\,\mathrm{d}t\right]^2\right)^{1/2}\right]^2\right.$$
$$+\sum_{n=1}^{m}\Delta t_n\sum_{K\in T_n} h_{n,K}^{2(k_{n,K}+1)}\|u(\cdot,t_n)\|^2_{H^{k_{n,K}+1}(K)}+\max_{0\leqslant n\leqslant m}\sum_{K\in T_n} h_{n,K}^{2(k_{n,K}+1)}\|u(\cdot,t_n)\|^2_{H^{k_{n,K}+1}(K)}$$
$$\left.+\left[\sum_{n=1}^{m}\Delta t_n\left\|\frac{\partial^2 u}{\partial\tau^2}\right\|_{L^1([t_{n-1},t_n];L^2(\Omega))}\right]^2+\left[\sum_{n=1}^{m}\|(R_n-R_{n-1})u_{n-1}\|\right]^2\right\}. \tag{4.6}$$

Proof. We follow the proof of Theorem 2.1 and techniques developed in [13]. Using the definition of the elliptic projection (2.8) and introducing the notation

$$e_n=U_n-R_nu_n,\qquad \hat{\bar{e}}_{n-1}=\hat{\bar{U}}_{n-1}-R_n\bar{u}_{n-1},$$
$$r_n=u_n-R_nu_n,\qquad \hat{\bar{r}}_{n-1}=\bar{u}_{n-1}-R_n\bar{u}_{n-1},$$

we have the following error equation for Algorithm 4.1:

$$\left(\phi\frac{e_n-\hat{\bar{e}}_{n-1}}{\Delta t_n},e_n\right)+(a_n\nabla e_n,\nabla e_n)+(c_ne_n,e_n)$$
$$=\left(\phi\frac{r_n-\hat{\bar{r}}_{n-1}}{\Delta t_n},e_n\right)+\left(\sqrt{\phi^2+|b_n|^2}\frac{\partial u}{\partial\tau}(t_n)-\phi\frac{u_n-\bar{u}_{n-1}}{\Delta t_n},e_n\right).$$

Using a change of variable technique, we can easily obtain $\|\hat{\bar{e}}_n\|^2_\phi\leqslant(1+C\Delta t_n)\|\hat{e}_n\|^2_\phi$. Thus the first term on the left-hand side of (4.7) can be estimated as

$$\left(\phi\frac{e_n-\hat{\bar{e}}_{n-1}}{\Delta t_n},e_n\right)\geqslant\frac{1}{2\Delta t_n}(\|e_n\|^2_\phi-\|\hat{\bar{e}}_{n-1}\|^2_\phi)\geqslant\frac{1}{2\Delta t_n}(\|e_n\|^2_\phi-\|\hat{e}_{n-1}\|^2_\phi)-C\|\hat{e}_{n-1}\|^2_\phi. \tag{4.8}$$

Applying Lemma 1 in [12] and (2.18) to the first term on the right-hand side of (4.7) we see that

$$\left(\phi\frac{r_n-\hat{\bar{r}}_{n-1}}{\Delta t_n},e_n\right)=\left(\phi\frac{r_n-\hat{r}_{n-1}}{\Delta t_n},e_n\right)+\left(\phi\frac{\hat{r}_{n-1}-\hat{\bar{r}}_{n-1}}{\Delta t_n},e_n\right)$$
$$\leqslant\|e_n\|_\phi\left\|\frac{r_n-\hat{r}_{n-1}}{\Delta t_n}\right\|_\phi+\|e_n\|_1\left\|\frac{\hat{r}_{n-1}-\hat{\bar{r}}_{n-1}}{\Delta t_n}\right\|_{-1}$$
$$\leqslant C\|e_n\|_\phi\Delta t_n^{-1}\left(\sum_{K\in T_n}\left[\int_{t_{n-1}}^{t_n} h_{n,K}^{k_{n,K}+1}\left\|\frac{\partial u}{\partial t}\right\|_{H^{k_{n,K}+1}(K)}\,\mathrm{d}t\right]^2\right)^{1/2}$$
$$+\frac{1}{2}(a_n\nabla e_n,\nabla e_n)+C\|\hat{r}_{n-1}\|^2. \tag{4.9}$$

The second term on the right-hand side of (4.7) can be estimated using Taylor's expansion along characteristics. Following the proof of Theorem 2.1, we can easily finish the rest of the proof. □

5. Dynamic mixed finite element schemes

Mixed finite element methods approximate the solution and its gradient simultaneously and transform the original second-order problem into a system of first-order equations. Define the flux $\sigma = -a\nabla u$ and the Sobolev space $H(\mathrm{div};\Omega) = \{v \in L^2(\Omega) : \nabla \cdot v \in L^2(\Omega)\}$, we have the mixed weak formulation for problem (2.1)–(2.3): find $\{u,\sigma\} \in L^2(\Omega) \times H(\mathrm{div};\Omega)$ such that

$$\left(\phi\frac{\partial u}{\partial t}, v\right) + (\nabla\cdot\sigma, v) - (b\cdot(a^{-1}\sigma), v) + (cu, v) = (f, v), \quad v \in L^2(\Omega), \tag{5.1}$$

$$(a^{-1}\sigma, \psi) - (u, \nabla\cdot\psi) = 0, \quad \forall\psi \in H(\mathrm{div};\Omega), \tag{5.2}$$

$$(u(\cdot,0), v) = (g, v), \quad \forall v \in L^2(\Omega). \tag{5.3}$$

Let $S_n \times X_n \subset L^2(\Omega)\times H(\mathrm{div};\Omega)$ be a mixed finite element space (satisfying the LBB condition; see [8]) at $t = t_n$. We assume that the following approximation property holds: for $n = 1,2,\ldots,N$,

$$\inf_{z\in S_n} \|v - z\| \leqslant C\left(\sum_{K\in T_n} h_{n,K}^{2(k_{n,K}+1)}\|v\|^2_{H^{k_{n,K}+1}(K)}\right)^{1/2}, \quad \forall v \in H^{k_n+1}(\Omega), \tag{5.4}$$

$$\inf_{w\in X_n} \|\psi - w\| \leqslant C\left(\sum_{K\in T_n} h_{n,K}^{2(k_{n,K}+1)}\|\psi\|^2_{H^{k_{n,K}+1}(K)}\right)^{1/2}, \quad \forall\psi \in H(\mathrm{div};\Omega)\cap H^{k_n+1}(\Omega), \tag{5.5}$$

where C is a constant independent of v, n, h_n, and k. Then a dynamic mixed finite element method reads:

Algorithm 5.1. *For $n = 1,2,\ldots,N$, compute first $\hat{U}_{n-1} \in S_n$ by (2.5) and then $\{U_n, W_n\} \in S_n \times X_n$ by*

$$\left(\phi\frac{U_n - \hat{U}_{n-1}}{\Delta t_n}, v\right) + (\nabla\cdot W_n, v) - (b_n\cdot(a_n^{-1}W_n), v) + (c_nU_n, v) = (f_n, v), \quad v \in S_n, \tag{5.6}$$

$$(a_n^{-1}W_n, w) - (U_n, \nabla\cdot w) = 0, \quad \forall w \in X_n. \tag{5.7}$$

Theorem 5.1. *Suppose that the solution u to problem (2.1)–(2.3) satisfies the regularity condition: $u, \partial u/\partial t, \nabla u \in L^\infty(H^{k+1}(\Omega)), \partial^2u/\partial t^2 \in L^1(L^2(\Omega))$. Let U_n be the solution of Algorithm 5.1. Then* *we have the error estimates for* $m = 1, 2, \ldots, N$,

$$\max_{1\leqslant n\leqslant m} \|u_n - U_n\|^2 + \sum_{n=1}^{m} \Delta t_n [(a_n^{-1}(\sigma_n - W_n), \sigma_n - W_n) + (c_n(u_n - U_n), u_n - U_n)]$$

$$\leqslant C \left\{ \|u_0 - U_0\|^2 + \left[\sum_{n=1}^{m} \left(\sum_{K\in T_n} \left[\int_{t_{n-1}}^{t_n} h_{n,K}^{k_{n,K}+1} \left\| \frac{\partial u}{\partial t} \right\|_{H^{k_{n,K}+1}(K)} \mathrm{d}t \right]^2 \right)^{1/2} \right]^2 \right.$$

$$+ \sum_{n=1}^{m} \Delta t_n \sum_{K\in T_n} h_{n,K}^{2(k_{n,K}+1)} \|\nabla u(\cdot, t_n)\|^2_{H^{k_{n,K}+1}(K)}$$

$$+ \max_{0\leqslant n\leqslant m} \sum_{K\in T_n} h_{n,K}^{2(k_{n,K}+1)} (\|u(\cdot, t_n)\|^2_{H^{k_n,K+1}(K)} + \|\nabla u(\cdot, t_n)\|_{H^{k_{n,K}+1}(K)})$$

$$\left. + \sum_{n=1}^{m} \left[\Delta t_n \int_{t_{n-1}}^{t_n} \left\| \frac{\partial^2 u}{\partial t^2} \right\| \mathrm{d}t \right]^2 + \left[\sum_{n=1}^{m} \|(R_n - R_{n-1}) u_{n-1}\| \right]^2 \right\}. \tag{5.8}$$

Proof. Introduce the elliptic projection $\{R_n u, R_n \sigma\}$ of $\{u, \sigma\}$: find $\{R_n u(\cdot, t), R_n \sigma(\cdot, t)\} \in S_n \times X_n$ for each $t \in [0, T]$ such that

$$(\nabla \cdot (R_n \sigma - \sigma), v) + (c_n(R_n u - u), v) = 0, \quad \forall v \in S_n, \tag{5.9}$$

$$(a_n^{-1} \cdot R_n \sigma, w) - (R_n u, \nabla \cdot w) = 0, \quad \forall w \in X_n. \tag{5.10}$$

and the notation

$$e_n = U_n - R_n u_n, \quad \hat{e}_{n-1} = \hat{U}_{n-1} - R_n u_{n-1},$$
$$r_n = u_n - R_n u_n, \quad \hat{r}_{n-1} = u_{n-1} - R_n u_{n-1},$$
$$\varepsilon_n = W_n - R_n \sigma_n, \quad \eta_n = \sigma_n - R_n \sigma_n.$$

Combining (5.6), (5.7), (5.1),(5.2), (5.9), and (5.10), we have

$$\left(\phi \frac{e_n - \hat{e}_{n-1}}{\Delta t_n}, v \right) + (\nabla \cdot \varepsilon_n, v) + (c_n e_n, v)$$
$$= \left(\phi \frac{r_n - \hat{r}_{n-1}}{\Delta t_n}, v \right) + \left(\phi \frac{\partial u}{\partial t}(t_n) - \phi \frac{u_n - u_{n-1}}{\Delta t_n}, v \right) + (b_n \cdot a_n^{-1}(W_n - \sigma_n), v), \quad \forall v \in S_n, \tag{5.11}$$

$$(a_n^{-1} \varepsilon_n, w) - (e_n, \nabla \cdot w) = 0, \quad \forall w \in X_n. \tag{5.12}$$

Taking $v = e_n$ in (5.11) and $w = \varepsilon_n$ in (5.12) and adding we obtain the error equation

$$\left(\phi \frac{e_n - \hat{e}_{n-1}}{\Delta t_n}, e_n \right) + (a_n^{-1} \varepsilon_n, \varepsilon_n) + (c_n e_n, e_n)$$
$$= \left(\phi \frac{r_n - \hat{r}_{n-1}}{\Delta t_n}, e_n \right) + \left(\phi \frac{\partial u}{\partial t}(t_n) - \phi \frac{u_n - u_{n-1}}{\Delta t_n}, e_n \right)$$
$$+ (b_n \cdot (a_n^{-1} \varepsilon_n), e_n) - (b_n \cdot (a_n^{-1} \eta_n), e_n). \tag{5.13}$$

Now the rest of the proof follows directly from Theorem 2.1. □

Crank–Nicolson and θ schemes can also be considered in a similar fashion. For nonlinear problems, our framework and error analysis presents no additional difficulty. Again, for other techniques such as [15–18,24,25], it is difficult to apply to general mixed finite element methods; such generalizations have not been seen in the literature.

6. Concluding remarks

We have analyzed a number of dynamic finite element methods for second-order linear and nonlinear parabolic equations using a unified framework. This framework enables us to study different finite element schemes and obtain unified convergence results. In particular, when the finite element space changes from time step to time step, the modified method of characteristics and mixed finite element methods are all treated in the same way as standard finite element methods for implicit Euler, Crank–Nicolson, and many other schemes. This has not been achieved by the theories provided in, for example, [15,13,6]. The convergence results obtained in this paper improve earlier ones and offer a clearer picture on the propagation of error due to the change of the finite element space. Our analysis also proves the convergence of some schemes which were not guaranteed by previous theory [15,13].

This paper has emphasized on the convergence theory of dynamic finite element methods and paid little attention to implementation issues. For example, where to apply fine grids and how to make grid refinement are very important problems in practice. There is a large literature on grid refinement strategies and here we just mention a few. Large gradient areas are usually the places where the solution develops steep layers or fronts. Thus predicting large gradient areas from the solution obtained at previous time step and making local grid refinement is one strategy [10]. In this respect, mixed finite element methods provide a more accurate prediction of the gradient and thus may be a good choice. One popular method among the engineering community, though, is to postprocess the approximate solution to obtain more accurate representation of the gradient. A posteriori error estimation is another way for doing adaptivity and local grid refinement [15,1,35,36]. When an approximate solution is obtained, the error between the approximate solution and the true solution can be estimated based on the information about the coefficients of the given partial differential equation and the approximate solution, which can be evaluated elementwise. Elements with large error are then subdivided into finer grids. Explicit a posteriori estimators can be computed directly from the finite element solution and the coefficients of the differential equation based on the residual equation, while implicit a posteriori estimators require solving local boundary value problems approximating the residual equation satisfied by the error.

No matter how local grid refinement and interpolation polynomial modification are made, our convergence theory states that the error between the exact solution and the approximate solution consists of three parts: a time finite difference discretization error, a spatial finite element discretization error, and an error term due to the projection of the approximated solution from old finite element spaces onto new finite element spaces. A good strategy to minimize the error would be that make grid refinement in a larger area to cover the local phenomena for several (maybe dozens of) time steps and that change the finite element space less frequently.

Numerical experiments have also shown that changing the grids at every time step or making grid refinement not according to the changing location of the local phenomena would affect the accuracy of the approximate solution. In [44], the author combined grid refinement and domain decomposition techniques to capture moving local phenomena for a model parabolic problem. Grid refinement was made only in subdomains that contain the local layer and coarse grid was applied in other subdomains. When the local layer moves, the domain was decomposed dynamically in such a way that the local layer was always contained in some subdomains, minimizing its intersection with interdomain boundaries, to improve accuracy.

Acknowledgements

The author would like to thank the referees for valuable suggestions.

References

[1] M. Ainsworth, J.T. Oden, A posteriori error estimation in finite element analysis, TICAM Report 96-19, Texas Institute for Computational and Applied Mathematics, University of Texas at Austin, May 1996.
[2] I. Babuska, B.Q. Guo, Approximation properties of the h–p versions of the finite element method, Comput. Methods Appl. Mech. Eng. 133 (1996) 319–346.
[3] I. Babuska, M. Suri, The p and h–p versions of the finite element method, basic principles and properties, SIAM Rev. 36 (1994) 578–632.
[4] M.J. Baines, An analysis of the moving finite element procedure, SIAM J. Numer. Anal. 28 (1991) 1323–1349.
[5] M.J. Baines, Moving Finite Elements, Oxford University Press, New York, 1994.
[6] R.E. Bank, R.F. Santos, Analysis of some moving space-time finite element methods, SIAM J. Numer. Anal. 30 (1993) 1–18.
[7] S.C. Brenner, L.R. Scott, The Mathematical Theory of Finite Element Methods, Springer, New York, 1994.
[8] F. Brezzi, M. Fortin, Mixed and Hybrid Finite Element Methods, Springer, New York, 1991.
[9] P.G. Ciarlet, The Finite Element Method for Elliptic Problems, North-Holland, Amsterdam, 1978.
[10] J.C. Diaz, R.E. Ewing, A.C. McDonald, L.M. Uhler, D.U. von Rosenberg, Self-adaptive local grid refinement for time-dependent two-dimensional simulations, in; R.H. Gallagher, D. Norrie, J.T. Oden, O.C. Zienkiewicz (Eds.), Finite Elements in Fluids, Vol. 6, Wiley, New york, 1985, pp. 279–290.
[11] J. Douglas Jr., T.F. Dupont, Galerkin methods for parabolic problems, SIAM J. Numer. Anal. 7 (1970) 575–626.
[12] J. Douglas Jr., T.F. Russell, Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with finite element or finite difference procedures, SIAM J. Numer. Anal. 19 (1982) 871–885.
[13] T.F. Dupont, Mesh modification for evolution equations, Math. Comput. 39 (1982) 85–108.
[14] K. Eriksson, D. Estep, P. Hansbo, C. Johnson, Introduction to adaptive methods for differential equations, Acta Numer. (1995) 105–158.
[15] K. Eriksson, C. Johnson, Adaptive finite element methods for parabolic problems I: a linear model problem, SIAM J. Numer. Anal. 28 (1991) 43–77.
[16] K. Eriksson, C. Johnson, Adaptive finite element methods for parabolic problems II: optimal error estimates in $L^\infty(L^2)$ and $L^\infty(L^\infty)$, SIAM J. Numer. Anal. 32 (1995) 706–740.
[17] K. Eriksson, C. Johnson, Adaptive finite element methods for parabolic problems IV: nonlinear problems, SIAM J. Numer. Anal. 32 (1995) 1729–1749.
[18] K. Eriksson, C. Johnson, Adaptive finite element methods for parabolic problems V: long-time integration, SIAM J. Numer. Anal. 32 (1995) 1750–1763.

[19] R.E. Ewing, T.F. Russell, M.F. Wheeler, Simulation of miscible displacement using mixed methods and a modified method of characteristics, SPE 12241, Seventh SPE Symposium on Reservoir Simulation, San Francisco, November 15–18, 1983.
[20] G.M. Hulbert, T.J.R. Hughes, Space-time finite element methods for second-order hyperbolic equations, Comput. Meth. Appl. Mech. Eng. 84 (1990) 327–348.
[21] J. Jaffre, C. Johnson, A. Szepessy, Convergence of the discontinuous Galerkin finite element method for hyperbolic conservation laws, Math. Models Methods Appl. Sci. 5 (1995) 367–386.
[22] P. Jamet, Galerkin-type approximations which are discontinuous in time for parabolic equations in a variable domain, SIAM J. Numer. Anal. 15 (1978) 912–928.
[23] P. Jamet, Stability and convergence of a generalized Crank–Nicolson scheme on a variable domain for the heat equation, SIAM J. Numer. Anal. 17 (1980) 530–539.
[24] C. Johnson, Numerical Solution of Partial Differential Equations by the Finite Element Method, Cambridge Univ. Press, Cambridge, 1987.
[25] C. Johnson, Discontinuous Galerkin finite element methods for second order hyperbolic problems, Comput. Meth. Appl. Mech. Eng. 107 (1993) 117–129.
[26] G.P. Liang, A finite element method with moving grid, Math. Numer. Sinica 7 (1985) 377–384.
[27] G.P. Liang, A finite element method of semidiscretization with moving grid, J. Comput. Math. 4 (1986) 86–96.
[28] G.P. Liang, Z.M. Chen, A full-discretization moving FEM with optimal convergence rate, Chinese J. Numer. Math. Appl. 12 (1990) 91–111.
[29] B. Lucier, A moving mesh numerical method for hyperbolic conservation laws, Math. Comput. 46 (1986) 59–69.
[30] K. Miller, Moving finite elements II, SIAM J. Numer. Anal. 18 (1981) 1033–1057.
[31] K. Miller, R.N. Miller, Moving finite elements I, SIAM J. Numer. Anal. 18 (1981) 1019–1032.
[32] H.X. Rui, A finite element method with moving grid for solving quasilinear pseudohyperbolic equations, Numer. Math. J. Chinese Univ. 10 (1988) 318–331.
[33] T.F. Russell, Finite elements with characteristic finite element method for a miscible displacement problem, SPE 10500, Proceedings of the Sixth SPE Symposium on Reservoir Simulation, Dallas, TX, 1982, pp. 123–135.
[34] E. Suli, C. Wilkins, Adaptive finite element methods for the damped wave equations, Report # 96/23, Oxford University Computing Lab, Oxford, 1996.
[35] R. Verfurth, A posteriori error estimates for nonlinear problems: $L^r(0,T;L^\rho(\Omega))$-error estimates for finite element discretizations of parabolic equations, Math. Comput. 67 (1998) 1335–1360.
[36] R. Verfurth, A posteriori error estimates for nonlinear problems: $L^r(0,T;W^{1,\rho}(\Omega))$-error estimates for finite element discretizations of parabolic equations, Numer. Methods Partial Differential Equations 14 (1998) 487–518.
[37] M.F. Wheeler, A priori L^2 error estimates for Galerkin approximation to parabolic partial differential equations, SIAM J. Numer. Anal. 10 (1973) 723–759.
[38] D.Q. Yang, The mixed finite element methods with moving grids for parabolic problems, Math. Numer. Sinica 10 (1988) 266–271.
[39] D.Q. Yang, Mixed methods with dynamic finite-element spaces for miscible displacement in porous media, J. Comput. Appl. Math. 30 (1990) 313–328.
[40] D.Q. Yang, A characteristic mixed method with dynamic finite element space for convection-dominated diffusion problems, J. Comput. Appl. Math. 43 (1992) 343–353.
[41] D.Q. Yang, Grid modification for the wave equation with attenuation, Numer. Math. 67 (1994) 391–401.
[42] D.Q. Yang, Grid modification for second order hyperbolic problems, Math. Comput. 64 (1995) 1495–1509.
[43] D.Q. Yang, Different domain decompositions at different times for capturing moving local phenomena, J. Comput. Appl. Math. 59 (1995) 39–48.
[44] D.Q. Yang, Dynamic domain decomposition and grid modification for parabolic problems, Comput. Math. Appl. 33 (1997) 89–103.
[45] D.Q. Yang, Numerical simulation of miscible displacement in porous media using an iterative perturbation method combined with a modified method of characteristics, Comput. Methods Appl. Mech. Eng. 162 (1998) 359–368.
[46] Y.R. Yuan, On finite element methods with moving mesh for 2-phase immiscible flow, Sci. Sinica, Ser. A 29 (1986) 785–799.
[47] Y.R. Yuan, On characteristic finite element methods with moving mesh for convection diffusion problems, Numer. Math. J. Chinese Univ. 8 (1986) 236–245.